Common Underlying Dynamics in an Emerging Market: From Minutes to Months

Renato Vicente† §, Charles M. de Toledo‡ , Vitor B.P. Leite ¶ and Nestor Caticha†

† Dep. de Física Geral, Instituto de Física, Universidade de São Paulo, Caixa Postal 66318, 05315-970 São Paulo - SP, Brazil
‡ BOVESPA - São Paulo Stock Exchange, R. XV de Novembro, 275, 01013-001 São Paulo - SP, Brazil
¶ Dep. de Física, IBILCE, Universidade Estadual Paulista, 15054-000 São José do Rio Preto - SP, Brazil

Abstract. We analyse a period spanning 35 years of activity in the São Paulo Stock Exchange Index (IBOVESPA) and show that the Heston model with stochastic volatility is capable of explaining price fluctuations for time scales ranging from 5 minutes to 100 days with a single set of parameters. We also show that the Heston model is inconsistent with the observed behavior of the volatility autocorrelation function. We deal with the latter inconsistency by introducing a slow time scale to the model. The fact that the price dynamics in a period of 35 years of macroeconomical unrest may be modeled by the same stochastic process is evidence for a general underlying microscopic market dynamics.

§ To whom correspondence should be addressed (rvicente@if.usp.br)

1. Introduction

In the last decades much attention has been devoted to the modeling of asset returns by the Finance community, the major drive behind this effort being the improvement of pricing techniques for derivative contracts [1]. The pricing problem is amenable to analytical solution for some stochastic volatility models [2] such as Hull-White [3], Stein-Stein [4] and Heston [5]. Despite differences in methods and emphasis, the cross fecundation between Economics and Physics, which dates back to the early nineteenth century (see [6] and [7]), has intensified recently [8]. Following the tradition of statistical physics, substantial effort has been made to find models capable of elucidating the basic mechanisms behind recurrent features of financial time series such as: returns aggregation (probability distributions at any time scale) [9, 10], volatility clustering [11], leverage effect (correlation between returns and volatilities) [12, 13], conditional correlations [14, 16].

Recently, a semi-analytical solution for the Fokker-Planck equation describing the distribution of log-returns in the Heston model has been proposed [9]. The authors were able to show a good agreement between the return distributions of a number of developed market stock indices and the model, for time scales spanning a wide interval ranging from 1 to 100 days.

In this paper we show strong evidence that the Heston model is capable of explaining the return distribution of the São Paulo Stock Exchange Index (IBOVESPA) in a period that span 35 years of political and economical unrest, with hyperinflation periods, currency crises and major regulatory changes. We also show that the Heston model can explain the diffusive process of IBOVESPA from minutes to months with a single set o parameters. However, we observe that the Heston model is inconsistent with the measured volatility autocorrelation function and we introduce an extension to the model along the lines discussed in [15] which exhibits the correct behavior.

2. The Heston Model

The Heston Model describes the dynamics of stock prices S_t as a geometric Brownian motion with volatility given by a Cox-Ingersoll-Ross (or Feller) mean-reverting dynamics. In the Itô differential form the model reads:

$$dS_t = S_t \, \mu_t dt \; + \; S_t \sqrt{v_t} \, dW_0(t) \tag{1}$$
$$dv_t = -\gamma \, [v_t - \theta] \, dt \; + \; \kappa \sqrt{v_t} \, dW_1(t),$$

where v_t is the volatility and dW_j are Wiener processes with

$$\langle dW_j(t) \rangle = 0, \qquad \langle dW_j(t) \, dW_k(\tilde{t}) \rangle = [\delta_{jk} + (1 - \delta_{jk})\rho] \; \delta(t - \tilde{t}) \, dt. \tag{2}$$

The term $\sqrt{v_t}$ avoids unphysical negative volatilities, θ is the macroeconomic long term volatility and μ_t represents a drift also at macroeconomic scales.

Interestingly, $v(t) \equiv \sum_{j=1}^{d} X_j^2(t)$ is a Feller process (1) if X_j are Ornstein-Uhlenbeck processes (OU) defined as [17]:

$$dX_j(t) = -\frac{b}{2} X_j(t)\, dt + \frac{a}{2}\, dW_j(t), \tag{3}$$

where dW_j describe d independent Wiener processes. This observation restricts the possible parameter values to $d \geq 2$ as the probability density of the volatility must vanish at zero to be consistent with the empirical data. To see how the volatility process in (1) emerges from OU processes we apply Itô's Lemma to get:

$$dv_t = -b\, dt \sum_{j}^{d} X_j^2 + a \sum_{j}^{d} X_j\, dW_j + \frac{a^2}{4} \sum_{j}^{d} dW_j^2. \tag{4}$$

Using the definition of v and the properties of the Wiener processes it follows that:

$$dv_t = \left[\frac{d}{4}\, a^2 - b v_t \right]\, dt + a\sqrt{v_t}\, dW. \tag{5}$$

The volatility process in (1) can be recovered by a few variable choices: $a = \kappa$, $b = \gamma$ and $\theta = \frac{d}{4}\frac{\kappa^2}{\gamma}$. Note that, given the dimension d, there are only two free parameters in (1). The OU processes in (3) may be regarded as the primary microscopic sources of volatility and the condition of non-vanishing volatility implies that $\alpha \equiv \frac{2\gamma\theta}{\kappa^2} > 1$.

3. The Fokker-Planck Equation Solution

We now outline the solution of the Fokker-Planck equation (FP) describing the distribution of log-returns proposed by Drăgulescu and Yakovenko [9].

As we are mainly concerned with price fluctuations, we simplify equation (1) by introducing log-returns in a window t as $r(t) = \ln(S(t)) - \ln(S(0))$. Using Ito's lemma and changing variables by making $x(t) = r(t) - \int_0^t d\tau \mu_\tau$ we obtain a detrended version of the return dynamics that reads:

$$dx = -\frac{v_t}{2}\, dt + \sqrt{v_t}\, dW_0. \tag{6}$$

The FP equation is, therefore, given by:

$$\frac{\partial P}{\partial t} = \gamma \frac{\partial}{\partial v}[(v - \theta)P] + \frac{1}{2}\frac{\partial}{\partial x}(vP) + \rho\kappa\frac{\partial^2}{\partial x \partial v}(vP) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(vP)$$

$$+ \frac{\kappa^2}{2}\frac{\partial^2}{\partial v^2}(vP), \tag{7}$$

that has to be solved for the following boundary condition $P(x, v, 0 \mid v_i) = \delta(x)\delta(v - v_i)$.

A Fourier transform in x followed by a Laplace transform in v leads to a partial differential equation of the first degree that can be solved by the method of characteristics [18]:

$$\left\{ \frac{\partial}{\partial t} + \left[\Gamma p_v + \frac{\kappa^2}{2}p_v^2 - \frac{(p_x^2 - ip_x)}{2} \right] \frac{\partial}{\partial p_v} \right\} Q = -\gamma\, p_v\, Q, \tag{8}$$

where $\Gamma \equiv i\rho\kappa p_x + \gamma$ and the boundary condition is $Q(p_x, p_v, 0 \mid v_i) = e^{-p_v v_i}$.

The unconditional distribution of log-returns can be obtained by inverting the Laplace and Fourier transforms and integrating first over the volatility v and then over the initial volatility v_i:

$$P_t(x) = \int_0^\infty dv_i \, P_v^*(v_i) \int_0^\infty dv \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} \, e^{ip_x x} \, Q(p_x, 0, t \mid v_i), \qquad (9)$$

where P_v^* is the stationary solution for the volatility distribution (see [9] for details).

Integrating (9) we finally get:

$$P_t(x) = \int_{-\infty}^{+\infty} \frac{dp_x}{2\pi} \, e^{ip_x x} \exp\left[\alpha\left(\frac{\Gamma t}{2} - \ln\cosh\left(\frac{\Omega t}{2}\right) - \Phi_t(p_x)\right)\right], \qquad (10)$$

where

$$\Omega \quad = \sqrt{\Gamma^2 + \kappa^2 \left(p_x^2 - ip_x\right)} \qquad (11)$$

$$\Phi_t(p_x) = \ln\left[1 + \frac{\Omega^2 - \Gamma^2 + 2\gamma\Gamma}{2\gamma\Omega} \tanh\left(\frac{\Omega t}{2}\right)\right]. \qquad (12)$$

4. Volatility Autocorrelation and Leverage Functions

The formal integral of (5) is given by:

$$v_t = (v_0 - \theta) \, e^{-\gamma_1 t} + \theta + \kappa \int_0^t dW_1(u) \, e^{-\gamma(t-u)} \sqrt{v_u}. \qquad (13)$$

A simple calculation gives the stationary autocorrelation function:

$$C(\tau \mid \gamma, \theta, \kappa) \equiv \lim_{t\to\infty} \frac{\langle v_t v_{t+\tau}\rangle - \langle v_t\rangle\langle v_{t+\tau}\rangle}{\theta^2} = \frac{e^{-\gamma\tau}}{\alpha}. \qquad (14)$$

The leverage function is [13]:

$$L(\tau \mid \gamma, \theta, \kappa, \rho) \equiv \lim_{t\to\infty} \frac{\langle dx_t \, (dx_{t+\tau})^2\rangle}{\langle (dx_{t+\tau})^2\rangle^2} = \rho\kappa \, H(\tau) \, G(\tau) \, e^{-\gamma\tau}, \qquad (15)$$

where dx_t is given by (6), $H(\tau)$ is the Heaviside step function and:

$$G(\tau) = \frac{\left\langle v_t \exp\left[\frac{\kappa}{2} \int_t^{t+\tau} dW_1(u) \, v_u^{-\frac{1}{2}}\right]\right\rangle}{\langle v_t\rangle^2}. \qquad (16)$$

To simplify the numerical calculations we employ throughout this paper the zeroth order appoximation $G(\tau) \approx G(0) = \theta^{-1}$. The approximation error increases with the time lag τ but is not critical to our conclusions.

In the next section we simultaneously fit (14), (15) and the model (10) to fluctuations of the BOVESPA index at a wide range of time scales.

5. Fitting IBOVESPA data

5.1. The data

Three data sets were used. IB1 consists of daily data from IBOVESPA inception on January, 1968 to December, 2002. IB1 was adjusted to take into account eleven divisions
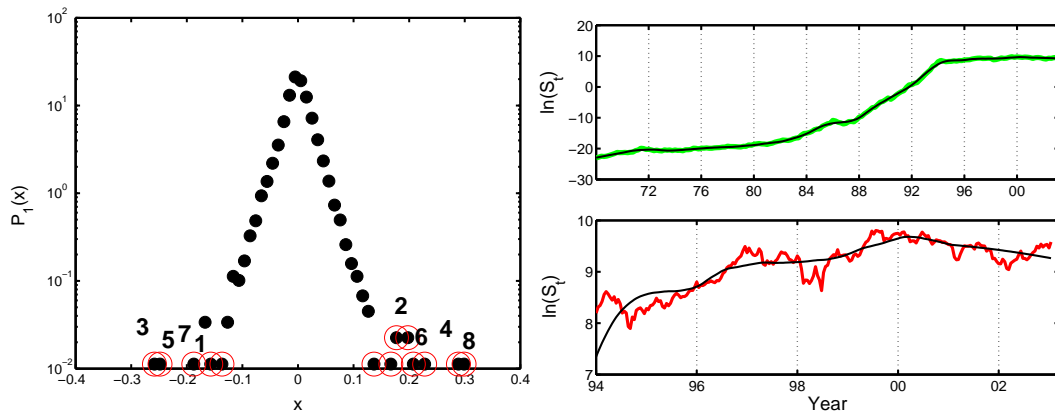
Figure 1. Left: Large deviations in daily log-returns with identifiable exogenous causes: New York Stock Market Crash (1); Economic Plans (2-4); Presidential crises (5), Recovery from Mexican, Asian and Russian crises (6); Asian and Russian crises (7) and Currency crises (8). Right: (Top) Macroeconomic trend in the log-prices. (Bottom) Trend and fluctuations in the last ten years.

by 10 introduced in the period for disclosure purposes [20]. IB2 consists of daily data from January, 2001 to August, 2003. IB3 consists of high-frequency data from March 3, 2001 to February 14, 2003 and from June 6, 2003 to August 26, 2003.

5.2. Trimming

Large deviations are explained both by endogenous dynamics and by exogenous shocks. An adequate microscopic model for the price formation process has to explain spontaneous large movements and the system response to external shocks. However, here we only intend to model the typical behavior at mesoscopic scales (price dynamics). In emerging markets, large movements with exogenous identifiable causes are frequent. In order to obtain a reliable fitting for the typical fluctuations we expunged from the data set large deviations connected to major structural changes, in Figure 1 we identify those externalities. For a brief historical account of recent Brazilian economy see [19].

5.3. Filtering the Macroeconomic Dynamics

In Equation 6 we extract the macroeconomic drift μ_t to focus on price fluctuations. This drift at long time scales reflects the effect of inflation, economic growth, business cycles and the riskless interest rate. To extract μ_t we have employed a low-pass Savitzky-Golay smoothing filter of degree two [21] with averaging over a four years moving window (*i.e.* one Brazilian presidential term). In Figure 1 we show $\int_0^t d\tau\,\mu_\tau$ for the whole data set (top) and the fluctuations around the trend in the last ten years (bottom). This choice of smoothing technique is heuristic and may introduce artifacts at time scales of lengths comparable to the moving windows (4 years in our case). We, therefore, restrict our analysis to a maximum period of 100 days and leave a thorough analysis of this aspect
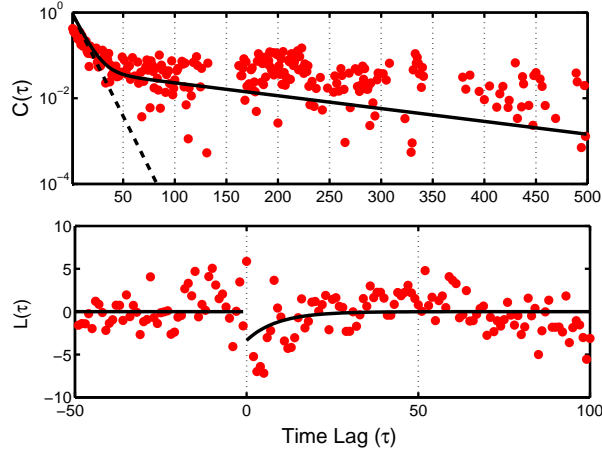
Figure 2. Top: Volatility autocorrelation: IB1 (circles), Heston model (dashed) and two time scales (solid line). Bottom: Leverage function: IB1 (circles) and approximate Heston model (see text, solid line).

to another occasion.

5.4. Fitting Parameters

After filtering the trend we have to fit four parameters: the long term mean volatility θ, the relaxation time for mean reversion $1/\gamma$, the volatility fluctuation scale κ and the correlation between price and volatility ρ. It became apparent in [10] that a simple least squares fit that takes into account only (10) yields parameters that are not uniquely defined. The task of finding parameters fitting all the main stilyzed facts simultaneously is not trivial. In this paper we do not intend to focus on parameter estimation, hence we adopt a heuristic procedure and do not provide rigorous error bars. These questions will be addressed somewhere else. We first estimate the long term mean volatility θ directly from the daily log-returns as:

$$\hat{\theta} = \frac{1}{N-1} \sum_{j=1}^{N} x_j^{(1)^2}, \tag{17}$$

where $x^{(1)}$ stands for daily detrended log-returns.

A proper fit must be also consistent with the non-vanishing volatilities constraint, $\frac{d}{2} > \frac{2\gamma\theta}{\kappa^2} \geq 1$. We, therefore, minimize a constraint free cost function and adjust κ to attain consistency with the constraint.

The following cost function is employed:

$$E(\gamma, \kappa, \rho) \equiv \frac{1}{I} \sum_{i=1}^{I} \left[1 - \frac{L^E(\tau_i)}{L(\tau_i \mid \gamma, \hat{\theta}, \kappa, \rho)} \right]^2 + \frac{1}{L} \sum_{l=1}^{L} \left[1 - \frac{C^E(\tau_l)}{C(\tau_l \mid \gamma, \hat{\theta}, \kappa)} \right]^2. \tag{18}$$
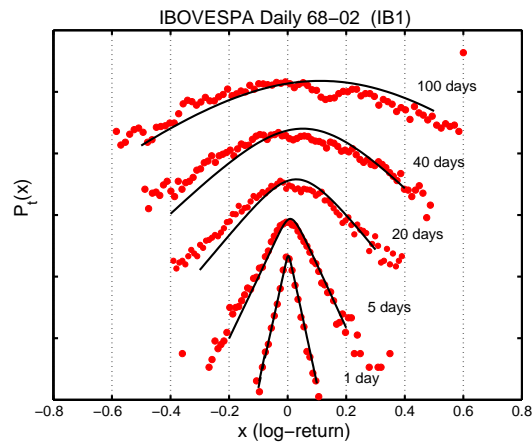
Figure 3. Probability densities of IBOVESPA returns from 1 to 100 days for the dataset IB1. The vertical scale is logarithmic and is multiplied by a constant for better visualization.

The empirical leverage function is [13]:

$$L^E(\tau) = \frac{\frac{1}{M}\sum_{t=1}^{M} x_t^{(1)} \left(x_{t+\tau}^{(1)}\right)^2}{\hat{\theta}^2},$$

(19)

where $x_t^{(1)} = \ln\left(\frac{S_t}{S_{t-1}}\right) - \mu_t$ and the empirical autocorrelation function is:

$$C^E(\tau) = \frac{\frac{1}{M}\sum_{t=1}^{M} \left(x_t^{(1)}\right)^2 \left(x_{t+\tau}^{(1)}\right)^2}{\hat{\theta}^2}.$$

(20)

We then employ the parameters obtained above to fit the probability density of returns and adjust the parameter ρ to the empirical data minimizing the mean squared error. We will discuss a more systematic approach to parameter estimation elsewhere as our main aim here is restricted to showing the description capability of the Heston model.

The parameters found are in the following table:

parameter	IB1	IB2-IB3
θ	$7.8 \times 10^{-4}\ days^{-1}$	$5.2 \times 10^{-4}\ days^{-1}$
$1/\gamma$	$9.0\ days$	$5.8\ days$
ρ	-0.20	-0.15
κ	$1.3 \times 10^{-2}\ days^{-1}$	$1.1 \times 10^{-2}\ days^{-1}$
d	2.03	2.98

Figure 2 shows IB1 data and curve fits. It is clear that, under the vanishing volatility constraint, a single relaxation time is not simultaneously consistent with leverage and autocorrelation functions. The best fit for the autocorrelation is shown as a dashed line. We deal with this inconsistency in the final sections of this paper by extending the Heston model to include two relaxation times.
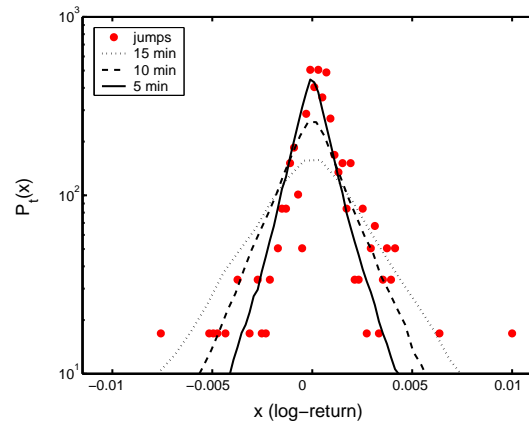
Figure 4. Probability densities of returns at 5, 10 and 15 minutes, overnight and lunch time jumps (dataset IB3).

5.5. Intraday Fluctuations

At first glance, it is not clear whether intraday and daily returns can be described by the same stochastic dynamics. Even less clear is whether aggregation from intraday to daily returns can be described by the same parameters. To verify this latter possibility we have to transform units by determining the effective duration in minutes of a business day T_{eff}. The dimensional parameters are, therefore, $\theta^{(ID)} = \theta/T_{eff}$, $\gamma^{(ID)} = \gamma/T_{eff}$ and $\kappa^{(ID)} = \kappa/T_{eff}$. In order to avoid the effects of non-stationarity we also use data from the same period (IB2 and IB3). From IB2 and IB3 we found the effective duration of a day to be about $T_{eff} = 540$ minutes. The São Paulo Stock Market opens daily at 11 a.m. and closes at 7:30 p.m. local time, resulting in exactly 510 minutes. The effective duration infered from data exceeds, therefore, in 30 minutes the real duration of a normal business day, an amount that is explained by the effect of overnight and lunchtime jumps. In Figure 4 we show a histogram of jumps compared to intraday returns at 5, 10 and 15 minutes. Considering both overnight and lunchtime jumps, between 10 to 30 minutes must be added to a normal business day, just the right amount to make T_{eff} compatible to the hypotheses of a single set of parameters describing time scales from minutes to months. In Figure 5 we show observations and data for IB2 and IB3 with parameters given by the table above. These findings are suported by similar results presented in [22] for high-frequency returns of individual stocks traded in NASDAQ and NYSE.

The datasets IB1 and IB3 yield $T_{eff} = 600$ minutes, the 90 minutes excess is explained by the well known non-stationarity of the parameters revealed in the table above, that shows that the mean variance in the 35 years period (IB1, $\theta = 7.8 \times 10^{-4}$ $days^{-1}$) is higher than the mean variance observed in the last two years (IB2, $\theta = 5.2 \times 10^{-4}$ $days^{-1}$). It is interesting to stress that in an emerging market like the Brazilian structural changes related to political events are frequent. In Figure 6 we try to put the the macroeconomic non-stationarity of the mean variance $\hat{\theta}$ in a historical perspective, identifying some important political events. The robust statistical behavior
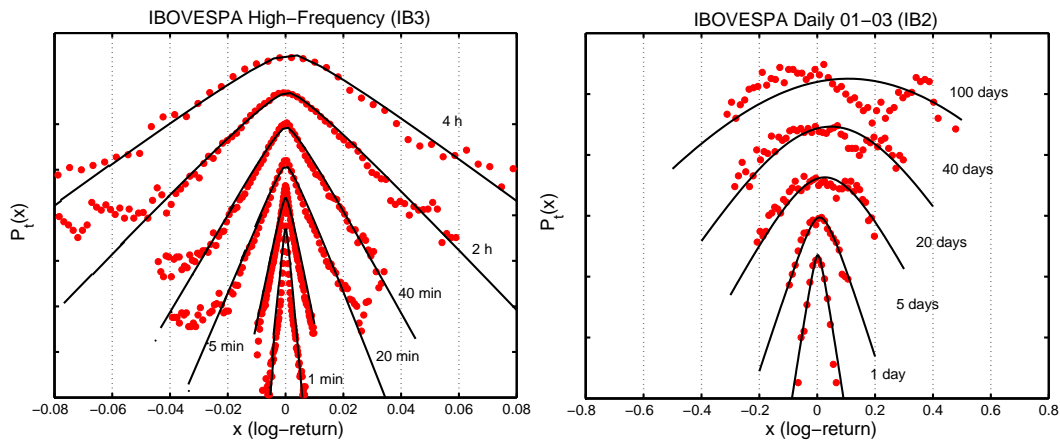
Figure 5. Left: Probability densities for intraday returns from 1 minute to 4 hours. Circles are empirical data from IB3 while lines correspond to the theory. Right: Returns from 1 to 100 days. Circles are empirical data from IB2. A single set of parameters (see Table) is used through all time scales considered. The vertical scale is logarithmic and is multiplied by a constant for better visualization.

here reported are particularly surprising when one considers the extremely abrupt structural changes that take place in an emerging market.

6. Multiple Time Scales

It is clear in Figure 2 that the Heston model is not capable of describing the autocorrelation function. To rectify this inconsistency we propose the inclusion of a second, slower, time scale into the dynamical model along the lines of [15]. We assume a stochastic dynamical model where the stochastic volatility reverts to a second stochastic volatility with much longer relaxation time $\gamma_1 \gg \gamma_2$. The new model reads:

$$dS = S(t)\mu dt \; + \; S(t)\sqrt{v(t)} \; dW_0(t) \tag{21}$$
$$dv = -\gamma_1 \left[v(t) - \theta\right] dt \; + \; \kappa_1 \sqrt{v(t)} \; dW_1(t),$$
$$d\theta = -\gamma_2 \left[\theta(t) - \theta_0\right] dt \; + \; \kappa_2 \sqrt{\theta(t)} \; dW_2(t),$$

where $\gamma_2/\gamma_1 << 1$ e dW_j are Wiener processes defined as:

$$\langle dW_j(t) \rangle \qquad = 0, \tag{22}$$
$$\langle dW_j(t) \, dW_k(\tilde{t}) \rangle = C_{jk}\delta(t - \tilde{t}) \; dt,$$

where C is a correlation matrix with $C_{jj} = 1$, $C_{12} = \rho$ and $C_{13} = C_{23} = 0$.

In this model the autocorrelation function acquires the following form:

$$C(\tau) = \frac{e^{-\gamma_1\tau}}{\alpha_1} + \frac{e^{-\gamma_2\tau}}{\alpha_2}, \tag{23}$$

with $\alpha_2 = \frac{2\gamma_2\bar{\theta}}{\kappa_2^2}$, where $\bar{\theta}$ stands for the average of θ given θ_0. In Figure 2 we show a fit of this autocorrelation function to the data, the new relaxation time is $\gamma_2^{-1} = 144.9$
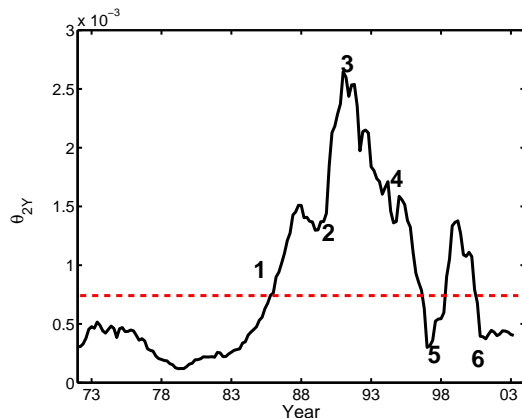
Figure 6. Historical perspective for the non-stationarity of the mean volatility calculated using the dataset IB1 and a two years wide moving window. The numbers identify the main political events as follows: military dictatorship ends after two decades (1), first presidential election (2), presidential impeachment (3), second election (4), reelection (5), third election (6).

days and $\kappa_2 = 1.0 \times 10^{-4}$ days^{-1}. It is possible to solve the Fokker-Planck equation for this extended model in the limit where $\alpha_2 \gg \alpha_1$. We will discuss this model in detail elsewhere [23].

7. Conclusions

We showed that the Heston model can reproduce the diffusion process of a 35 years long time series of IBOVESPA returns at a wide range of time scales from days to months. We also show that the Heston model can explain the aggregation of returns from minutes to months with a single set o parameters that change in the macroeconomic scales. However, the Heston model, with a single relaxation time, is not capable of explaining the behavior of the autocorrelation function. We, therefore, proposed an extended version of the model with the addition of a slow stochastic dynamics that yields an autocorrelation function consistent with the observations.

It is surprising that a single non-trivial stochastic model may be capable of explaining the long-term statistical behavior of both developed and emerging markets, despite the known instability and high susceptibility to externalities of the latter. We believe that this robust non-trivial statistical behavior is evidence for more basic mechanisms acting in the market microstructure. Perhaps the search for underlying common mechanisms (or laws) that can explain empirical data is the main contribution of Physics to Economics. This contribution might be particularly useful to the field of Econometrics in which a common view is that a theory built from data 'should be evaluated in terms of the quality of the decisions that are made based on the theory' [24]. Clearly, these two approaches should not be considered as mutually exclusive.

Acknowledgments

[1] J.P. Fouque, G. Papanicolaou, K.R. Sircar, Derivatives in Financial Markets with Stochastic Volatility, Cambridge University Press, Cambridge, 2000.

[2] E. Gysels, A.C. Harvey, E. Renault, Stochastic Volatility, in Handbook of Statistics 14, eds. G.S. Madalla, C.R. Rao, Elsevier, Amsterdan, 1996.

[3] J. Hull, A. White, Journal of Finance 42 (1987) 281.

[4] E.M. Stein, J.C. Stein, Review of Financial Studies 4 (1991) 727.

[5] S.L. Heston, Review of Financial Studies 6 (1993) 327.

[6] B.M. Roehner, Patterns of Speculation, Cambridge University Press, Cambridge, 2002.

[7] P. Mirowski, More Heat Than Light, Cambridge University Press, Cambrigde, 1989.

[8] R.N. Mantegna, H.E. Stanley, An Introduction to Econophysics, Cambridge University Press, Cambridge, 2000.

[9] A.A.Drăgulescu, V.M. Yakovenko, Quantitative Finance 2 (2002) 443.

[10] A.C. Silva, V.M. Yakovenko, Physica A 324 (2003) 303.

[11] R.F. Engle, A.J. Patton, Quantitative Finance 1 (2001) 237.

[12] J.-P. Bouchaud, A. Matacz, M. Potters, Physical Review Letters 87 (2001) 228701-1.

[13] J. Perelló, J. Masoliver, Physical Review E 67 (2003) 037102.

[14] M. Boguñá, J. Masoliver, *Preprint* cond-mat/0310217 (2003).

[15] J. Perelló, J. Masoliver, J.-P. Bouchaud, *Preprint* cond-mat/0302095 (2003).

[16] B. LeBaron, Journal of Applied Econometrics, 7 (1992) S137.

[17] S. Shreve, Lectures on Stochastic Calculus and Finance.

[18] D. Zwillinger, *Handbook of Differential Equations*, Second Edition, Academic Press (1992).

[19] R.L. Costa, G.L. Vasconcelos, Physica A 329 (2003) 231.

[20] BOVESPA index Webpage, *http://www.bovespa.com.br/indbovi.shtml*.

[21] W.H.Press et al., *Numerical Recipes in C: The Art of Scientific Computing*, Second Edition, Cambridge University Press (2002).

[22] A.C. Silva, R. E. Prange, V.M. Yakovenko, *Preprint* cond-mat/0401225 (2004).

[23] R. Vicente, V.B.P. Leite, N. Caticha, *in preparation*.

[24] C.W.J. Granger, *Empirical Modeling in Economics*, First Edition, Cambridge University Press (1999).